

# THE MIXING TIME OF THE NEWMAN–WATTS SMALL WORLD

LOUIGI ADDARIO-BERRY AND TAO LEI

**ABSTRACT.** “Small worlds” are large systems in which any given node has only a few connections to other points, but possessing the property that all pairs of points are connected by a short path, typically logarithmic in the number of nodes. The use of random walks for sampling a uniform element from a large state space is by now a classical technique; to prove that such a technique works for a given network, a bound on the mixing time is required. However, little detailed information is known about the behaviour of random walks on small-world networks, though many predictions can be found in the physics literature. The principal contribution of this paper is to show that for a famous small-world random graph model known as the Newman–Watts small world, the mixing time is of order  $\log^2 n$ . This confirms a prediction of Richard Durrett, who proved a lower bound of order  $\log^2 n$  and an upper bound of order  $\log^3 n$ .

## 1. INTRODUCTION

The small-world phenomenon is a catchy name for an important physical phenomenon that shows up throughout the physical, biological and social sciences. In brief, the term applies to large, locally sparse systems (usually, possessing only a bounded number of connections from any given point) which nonetheless exhibit good long-range connectivity in the sense that there are short paths between all points in the system. The Erdős–Rényi random graph  $G_{n,p}$ , is perhaps the most mathematically famous model possessing small-world behaviour: when  $p = c/n$  for  $c > 1$  fixed, the average vertex degree is  $c$ , and the diameter of the largest connected component is [18]

$$\frac{\log n}{\log c} + 2 \frac{\log n}{\log(1/c^*)} + O_p(1),$$

where  $c^* < 1$  satisfies  $ce^{-c} = c^*e^{-c^*}$  and  $O_p(1)$  denotes a random amount that remains bounded in probability as  $n \rightarrow \infty$ .

The Erdős–Rényi random graph is unsatisfactory as a small-world model in two ways: first, the network does not satisfy full connectivity (a constant proportion of vertices lie outside of the giant component); second, the graph is locally tree-like – for any fixed  $k$ , the probability that there is a cycle of length at most  $k$  through a randomly chosen node is  $o(1)$ . In real-world networks showing small world behaviour (social or business networks, gene regulatory networks, networks for modelling infectious disease spread, scientific collaboration networks, and many others – the book [16] contains many interesting examples), full or almost-full connectivity is standard, and short cycles are plentiful. Several connected models have been proposed which in some respects capture the desired local structure as well as small-world behaviour, notably the Bollobás–Chung [1], Watts–Strogatz [20], and Newman–Watts [14, 15] models. These models are closely related – all are based on adding sparse, long range connections to a connected “base network” which is essentially a cycle.

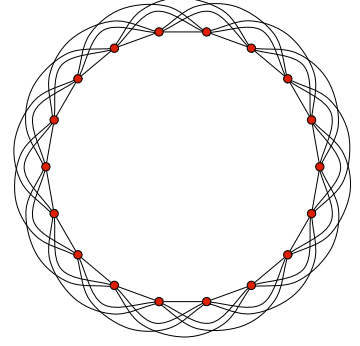
*Date:* January 19, 2012.

*2000 Mathematics Subject Classification.* 60C05.

*Key words and phrases.* Random graph, small world, mixing time, rapid mixing.

Understanding the behaviour of random walks on small-world networks remains, in general, a challenging open problem. Numerical and non-rigorous results for return probabilities [10], relaxation times [17], spectral properties [5], hitting times [3, 9], and diffusivity [7] appear in the physics literature, but few rigorous results are known. In [4], Durrett considers the Newman–Watts small world, proving lower and upper bounds on the mixing time of order  $\log^2 n$  and  $\log^3 n$ , respectively, and suggests that the lower bound should in fact be correct. The principle contribution of this paper is to confirm Durrett’s prediction. (Theorem 1, below).

To define the Newman–Watts small world, first fix integers  $n \geq k \geq 1$ . The  $(n, k)$ -ring  $R_{n,k}$  is the graph with vertex set  $[n] = \{1, \dots, n\}$  and edge set  $\{\{i, j\} : i + 1 \leq j \leq i + k\}$ , where addition is interpreted modulo  $n$ . (A picture of  $R_{18,3}$  appears to the right.) In particular, an  $(n, 1)$ -ring is a cycle of length  $n$ , and whenever  $n > 2k$  the  $(n, k)$  ring is regular of degree  $2k$ . For  $0 < p < 1$ , the  $(n, k, p)$  Newman–Watts small world  $H_{n,k,p}$  is the random graph obtained from the  $(n, k)$ -ring by independently replacing each non-edge of the  $(n, k)$ -ring by an edge with probability  $p$ . We write  $H$  as shorthand for  $H_{n,k,p}$  whenever the parameters are clear from context.



Given a (finite, simple) graph  $G = (V, E)$ , by a *lazy simple random walk on  $G$*  we mean a random walk that at each step stays still with probability  $1/2$ , and otherwise moves to a uniformly random neighbour. In other words, this is a Markov chain with state space  $V$  and transition probabilities

$$p_{x,y} = \begin{cases} \frac{1}{2} & \text{if } x = y \\ \frac{1}{2d_G(x)} & \text{if } y \in N_G(x) \\ 0 & \text{otherwise} \end{cases}$$

where  $d_G(x)$  denotes the number of neighbours of  $x$  in  $G$  and  $N_G(x)$  denotes the collection of neighbours of  $x$  in  $G$ . (More generally, we shall say a chain is *lazy* if  $p_{x,x} \geq 1/2$  for all  $x$  in the state space.) If  $G$  is connected then this Markov chain has a unique stationary distribution  $\pi$  given by  $\pi(x) = d_G(x)/2|E|$ .

Given two probability distributions  $\mu, \nu$  on  $V$ , we define the *total variation distance* between  $\mu$  and  $\nu$  as

$$\|\mu - \nu\|_{\text{TV}} = \sup_{S \subset V} |\mu(S) - \nu(S)| = \frac{1}{2} \sum_{v \in V} |\mu(v) - \nu(v)|,$$

where  $\mu(S) = \sum_{v \in S} \mu(v)$ . Now let  $(X_k)_{k \geq 0}$  be a lazy simple random walk on  $G$ , and write  $\mu_{k,x}$  for the distribution of  $X_k$  when the walk is started from  $x$ ; formally, for all  $y \in G$ ,  $\mu_{k,x}(y) = \mathbf{P}(X_k = y | X_0 = x)$ . The *mixing time* of the lazy simple random walk on  $G$  is defined as

$$\tau_{\text{MIX}}(G) = \max_x \min\{k : \|\mu_{k,x} - \pi\|_{\text{TV}} \leq 1/4\}.$$

(There are many different notions of mixing time, many of which are known to be equivalent up to constant factors – the book [11], and the survey [12] are both excellent references.) We may now formally state our main result.

**Theorem 1.** *Fix  $c > 0$ , let  $p = c/n$ , and let  $H$  be an  $(n, k, p)$  Newman–Watts small world. Then there is  $C_0 > 0$  depending only on  $c$  and  $k$  such that with probability at least  $1 - O(n^{-3})$ ,*

$$C_0^{-1} \log^2 n \leq \tau_{\text{MIX}}(H) \leq C_0 \log^2 n.$$

Furthermore,  $\mathbf{E}[\tau_{\text{MIX}}(H)] \leq C_0(\log^2 n + 1)$ .

The expectation bound in Theorem 1 follows easily from the probability bound. Indeed, given a finite reversible lazy chain  $X = (X_t, t \geq 0)$  with state space  $\Omega$ , for  $x \in \Omega$  write  $\tau_x = \min\{t \geq 0 : X_t = x\}$  for the hitting time of state  $x$ . Then  $\tau_{\text{MIX}} \leq 2 \max_{x \in \Omega} \mathbb{E}_\pi(\tau_x) + 1$ , where  $\mathbb{E}_\pi$  denotes expectation starting from stationarity (see, e.g., [11], Theorem 10.14 (ii)). If  $X$  is a lazy simple random walk on a connected graph  $G = (V, E)$  with  $|V| = n$ , then  $\max_{x \in V} \mathbb{E}_\pi(\tau_x) \leq (4/27 + o(1))n^3$  (see [2]). Assuming the probability bound of Theorem 1 and applying the two preceding facts we obtain that

$$\mathbb{E}[\tau_{\text{MIX}}(H)] \leq C_0 \log^2 n + \frac{8}{27} n^3 (1 + o(1)) \cdot \mathbf{P}(\tau_{\text{MIX}}(H) > C_0 \log^2 n) \leq C_0 (\log^2 n + 1),$$

assuming  $C_0$  is chosen large enough.

The lower bound of Theorem 1 is also straightforward, and we now provide its proof. For  $n$  sufficiently large, given  $v \in [n]$  and  $\ell \in \mathbb{N}$  the probability that all vertices  $w$  of  $H_{n,k,p}$  with  $w \in [v - \ell, v + \ell] \bmod n$  have degree exactly  $2k$ , is greater than

$$(1 - p)^{n(2\ell+1)} \geq e^{-2c(2\ell+1)},$$

the inequality holding since  $(1 - c/n)^n > e^{-2c}$  for  $n$  large. Taking  $\alpha = 1/(8c)$ , it follows easily that with probability at least  $1 - O(n^{-3})$  there is  $v \in [n]$  such that all vertices  $w$  with  $w \in [v - \alpha \log n, v + \alpha \log n] \bmod n$  have degree exactly  $2k$ . Furthermore, the random walker started from such a vertex  $v$  will with high probability take time of order  $\log^2 n$  before first visiting a vertex in the complement of  $[v - \alpha \log n, v + \alpha \log n] \bmod n$ . Finally, under  $\pi$ , the set  $[v - \alpha \log n, v + \alpha \log n] \bmod n$  has measure tending to zero with  $n$ , and it thus follows from the definition of  $\tau_{\text{MIX}}$  that the mixing time is of order at least  $\log^2 n$  whenever such a vertex  $v$  exists.

Having taken care of the expectation upper bound and lower bound in probability from Theorem 1, the remainder of the paper is now devoted to proving that with probability at least  $1 - O(n^{-3})$ ,  $\tau_{\text{MIX}}(H) = O(\log^2 n)$ . In Section 2 we explain a conductance-based mixing time bound of Fountoulakis and Reed [6] which will form the basis of our approach. The Fountoulakis–Reed bound requires control on the edge expansion of connected subgraphs of  $H_{n,k,p}$ . To this end, in Section 3 we bound the expected *number* of connected subgraphs of  $H_{n,k,p}$  of size  $j$ , for each  $1 \leq j \leq n$ ; our probability bounds on the edge expansion of such subgraphs follow in Section 4. Finally, in Section 5 we finish the proof of Theorem 1. The proof is more straightforward when  $c$  is large; in order to get the key ideas across cleanly we accordingly handle the large- $c$  and small- $c$  cases separately.

**1.1. Notation.** Given a graph  $G$ , write  $V(G)$  for the set of vertices of  $G$ , and  $E(G)$  for the set of edges of  $G$ . Also, given  $S \subset V(G)$ , write  $G[S]$  for the subgraph of  $G$  induced by  $S$ . We say  $S$  is connected if  $G[S]$  is connected. Finally, given a formal power series  $F(z)$ , we write  $[z^j]F(z)$  to mean the coefficient of  $z^j$  in  $F(z)$ , so if  $F(z) = \sum_{k \geq 0} a_k z^k$  then  $[z^j]F(z) = a_j$ .

## 2. MIXING TIME VIA CONDUCTANCE BOUNDS

A range of techniques are known for bounding mixing times ([13] is a recent survey of the available approaches), many of which are tailor-made to give sharp bounds for particular families of chains. One particularly fruitful family of techniques is based on bounding the *conductance* of the underlying graph, a function which encodes the presence of bottlenecks at all scales. The precise bound we shall use is due to Fountoulakis and Reed [6]. Given sets  $S, T \subset V$ , write  $E(S, T) = E_G(S, T)$  for the set of edges of  $G$  with one endpoint in  $S$  and the

other in  $T$ , and write  $e(S, T) = |E(S, T)|$ . Also, given  $S \subset V$  write  $e(S) = \sum_{v \in S} d_G(v)$ . The *conductance* of  $S$ , written  $\Phi(S)$ , is given by

$$\Phi(S) = \frac{e(S, S^c)}{e(S)}.$$

For  $0 \leq x \leq 1/2$ , write

$$\Phi(x) = \min_{\substack{S \text{ connected} \\ x|E| \leq e(S) \leq 2x|E|}} \Phi(S);$$

one can think of  $\Phi(x)$  the (worst case) *connected conductance of  $G$  at scale  $x$* . We will use the following theorem, a specialization of the main result from [6].

**Theorem 2** ([6]). *There exists a universal constant  $C > 0$  so that for any connected graph  $G$ ,*

$$\tau_{\text{MIX}}(G) \leq C \sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi^{-2}(2^{-i}).$$

With this theorem at hand, proving mixing time bounds boils down to understanding what sorts of bottlenecks can exist in  $G$ . For the  $(n, k, p)$  Newman–Watts small world with  $p = c/n$ , it is not hard to see that small sets can have poor conductance. Indeed, in the introduction we observed that with high probability the ring  $R_{n,k}$  will contain connected sets  $S$  with  $\Theta(\log n)$  nodes, to which no edges are added in  $H_{n,k,p}$ . Such a set  $S$  will have  $e(S, S^c) < k^2$  and so will have conductance  $\Phi(S) = O(1/\log n)$ . It follows that in this case the best possible mixing time bound one can hope to prove using Theorem 2 is of order  $\log^2 n$ .

We will prove the upper bound in the probability bound of Theorem 1 by showing that there are constants  $\epsilon > 0$ ,  $C_0 > 0$  such that with high probability, whenever  $|S| \geq C_0 \log n$ , we have  $\Phi(S) \geq \epsilon$ . To accomplish this using Theorem 2, we will need control on the likely number of connected subgraphs of  $H_{n,k,p}$  of size  $s$ , for all  $s \geq \log n$ . In the next section, we bound the expected number of such subgraphs using Lagrange inversion and comparison with a branching process.

### 3. COUNTING CONNECTED SUBGRAPHS

Fix  $c > 0$  and  $k \geq 1$ , let  $p = c/n$ , and let  $H = H_{n,k,p}$  be a Newman–Watts small world. Let  $v \in [n]$ , write  $B_{j,v} = B_{j,v}(H)$  for the set of all  $S \subset [n]$  containing  $v$  with  $|S| = j$  such that  $H[S]$  is connected, and let  $B_j = \bigcup_{v \in [n]} B_{j,v}$ . Our aim in this section is to establish the following proposition.

**Proposition 3.** *For any positive integer  $j$  and any  $v \in [n]$ ,  $\mathbf{E}|B_{j,v}| \leq (4(c + 2k))^j$ , and  $\mathbf{E}|B_j| \leq n(4(c + 2k))^j$ .*

We will prove Proposition 3 by comparison with a Galton–Watson process. Recall that a Galton–Watson process can be described as follows. An initial individual – the progenitor – has a random number  $Z_1$  of children, where  $Z_1$  is some non-negative, integer-valued random variable. The distribution of  $Z_1$  is called the *offspring distribution*. Each child of the progenitor reproduces independently according to the offspring distribution, and this process continues recursively. The family tree of a Galton–Watson process is called a Galton–Watson tree, and is rooted at the progenitor.

The number of neighbours of a vertex in  $H_{n,k,p}$  is distributed as  $\text{Bin}(n - 2k - 1, p) + 2k$ . From this it is easily seen that  $|B_{j,v}|$  is stochastically dominated by the number of subtrees of

size  $j$  containing the root in a Galton–Watson tree with offspring distribution  $\text{Bin}(n - 2k - 1, p) + 2k$ . To bound the expectations of the latter random variables, we will first encode these expectations as the coefficients of a generating function, then use the Lagrange inversion formula ([19], Theorem 5.4.2), which we now recall. This approach was suggested to us by Omer Angel in a mathoverflow comment (<http://mathoverflow.net/questions/66595/>); we thank him for the suggestion.

**Theorem 4** (Lagrange inversion formula). *If  $G(x)$  is a formal power series and  $f(x) = xG(f(x))$ , then*

$$n[x^n]f(x)^k = k[x^{n-k}]G(x)^n.$$

Fix a non-negative, integer-valued random variable  $B$ , and for  $m \geq 0$  write  $p_m = \mathbf{P}(B = m)$ . Given a Galton–Watson tree  $\mathcal{T}$  with offspring distribution  $B$ , let  $\mu_j = \mu_j(B)$  denote the expected number of subtrees of  $\mathcal{T}$  containing the root of  $\mathcal{T}$  and having exactly  $j$  vertices (so  $\mu_0 = 0$ ). Also, write

$$q_j = \sum_{m \geq j} p_m(m)_j,$$

where  $(m)_j = m!/(m-j)!$  is the falling factorial. Note that  $q_j$  is the expected number of ways to choose and order precisely  $j$  children of the root in  $\mathcal{T}$ . Let  $F(z) = \sum_{j=0}^{\infty} \mu_j z^j$  and  $Q(z) = \sum_{j=0}^{\infty} q_j z^j$  be the generating functions of  $\mu_j$  and  $q_j$  respectively, viewed as formal power series.

**Lemma 5.**  $F(z) = zQ(F(z))$

*Proof.* We have

$$\begin{aligned} Q(F(z)) &= \sum_{j \geq 0} q_j \left( \sum_{r \geq 1} \mu_r z^r \right)^j \\ &= \sum_{j \geq 0} q_j \left( \sum_{r \geq j} \sum_{\substack{r_1 + \dots + r_j = r \\ r_1, \dots, r_j \in \mathbb{N}^+}} z^r \mu_{r_1} \dots \mu_{r_j} \right) \\ &= \frac{1}{z} \sum_{r \geq 0} z^{r+1} \sum_{j \leq r} q_j \left( \sum_{\substack{r_1 + \dots + r_j = r \\ r_1, \dots, r_j \in \mathbb{N}^+}} \mu_{r_1} \dots \mu_{r_j} \right) \end{aligned} \tag{1}$$

The  $r$ 'th term in the outer sum in (1) encodes subtrees of  $\mathcal{T}$  with  $r+1$  vertices that contain the root, as follows. First specify the degree  $j$  of the root of the tree  $T$  to be embedded. Then choose which  $j$  children of the root of  $\mathcal{T}$  will form part of the embedding, and the order in which the children of the root of  $T$  will be mapped to these nodes (there are  $q_j$  ways to do this on average). Next, choose the sizes  $r_1, \dots, r_j$  of the subtrees of the children of the root in the embedded tree; finally, embed each such subtree in the respective subtree of  $\mathcal{T}$ ; on average, there are  $\mu_{r_i}$  ways to do this. It follows that

$$\begin{aligned} (1) &= \frac{1}{z} \sum_{r \geq 0} z^{r+1} \mu_{r+1} \\ &= \frac{1}{z} F(z), \end{aligned} \tag{2}$$

which proves the lemma. (We remark that verifying that (1) and (2) are equal can be done purely formally; however, we find the preceding explanation more instructive.)  $\square$

**Lemma 6.** *Fix  $C > 0$ . if  $q_j \leq C^j$  for all  $j \geq 0$  then  $\mu_j \leq \frac{1}{j} \binom{2j-2}{j-1} C^{j-1} < (4C)^{j-1}$  for all  $j \geq 1$ .*

*Proof.* By Lemma 5 and Theorem 4, we have  $j[z^j]F(z)^k = k[z^{j-k}]Q(z)^j$ . In particular, taking  $k = 1$ , we have  $\mu_j = [z^j]F(z) = \frac{1}{j}[z^{j-1}]Q(z)^j$ . Now,

$$(Q(z))^j = \left( \sum_{l \geq 0} q_l z^l \right)^j = \sum_{r \geq 0} \left( \sum_{\substack{l_1 + \dots + l_j = r \\ l_1, \dots, l_j \in \mathbb{N}}} q_{l_1} q_{l_2} \dots q_{l_j} \right) z^r.$$

Therefore  $[z^{j-1}]Q(z)^j = \sum_{\substack{l_1 + \dots + l_j = j-1 \\ l_1, \dots, l_j \in \mathbb{N}}} q_{l_1} q_{l_2} \dots q_{l_j}$ . Each summand  $q_{l_1} \dots q_{l_j}$  is at most  $C^{j-1}$  by assumption. There are  $\binom{2j-2}{j-1}$  nonnegative integer solutions to the equation  $l_1 + \dots + l_j = j-1$ , so we obtain that  $[z^{j-1}]Q(z)^j \leq \binom{2j-2}{j-1} \cdot C^{j-1}$ . The result follows.  $\square$

The next lemma controls the growth of  $q_j$  for some important special offspring distributions, which allows us to use Lemma 6 to prove Proposition 3.

**Lemma 7.** *If  $B$  is Poisson( $c$ ) distributed then  $q_j = c^j$  for all  $j$ . Also, if  $B$  is  $\text{Bin}(n, c/n)$  distributed, then for all  $j \geq 0$ ,  $q_j \leq c^j$ . Finally, if  $B - \ell$  is  $\text{Bin}(n, c/n)$  distributed for some fixed  $\ell \geq 0$ , then for all  $j \geq 0$ ,  $q_j \leq (c + \ell)^j$ .*

*Proof.* If  $B$  is Poisson with mean  $c$ , then

$$q_j = \sum_{m \geq j} p_m \frac{m!}{(m-j)!} = e^{-c} c^j \left( 1 + \frac{c}{1!} + \frac{c^2}{2!} + \dots \right) = e^{-c} c^j e^c = c^j$$

If  $B \stackrel{d}{=} \text{Bin}(n, c/n)$  then

$$\begin{aligned} q_j &= \sum_{j \leq m \leq n} \binom{n}{m} \left( \frac{c}{n} \right)^m \left( 1 - \frac{c}{n} \right)^{n-m} \frac{m!}{(m-j)!} \\ &= \left( \frac{c}{n} \right)^j \cdot n(n-1) \dots (n-j+1) \sum_{j \leq m \leq n} \binom{n-j}{n-m} \left( \frac{c}{n} \right)^{m-j} \left( 1 - \frac{c}{n} \right)^{n-m} \\ &\leq c^j, \end{aligned}$$

Finally, if  $B \stackrel{d}{=} \text{Bin}(n, c/n) + \ell$  then we consider three Galton-Watson trees  $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$  with offspring distributions  $B_1 \equiv \ell, B_2 \stackrel{d}{=} \text{Bin}(n, c/n)$  and  $B_3 \stackrel{d}{=} \text{Bin}(n, c/n) + \ell$  respectively. For  $i = 1, 2, 3$  write  $q_j^{(i)} = q_j(B_i)$ . Since  $q_j^{(1)}$  is the expected number of ways to choose and order precisely  $j$  children of the root in  $\mathcal{T}_1$ , we have  $q_j^{(1)} = (\ell)_j \leq \ell^j$  and by the previous argument we know that  $q_j^{(2)} \leq c^j$ . Finally, since  $q_j^{(3)}$  is the expected number of ways to choose and order

precisely  $j$  children of the root in  $\mathcal{T}_3$ , by independence we have

$$\begin{aligned} q_j^{(3)} &= \sum_{s=0}^j \binom{j}{s} q_s^{(1)} q_{j-s}^{(2)} \\ &\leq \sum_{s=0}^j \binom{j}{s} l^s c^{j-s} \\ &= (l + c)^j. \end{aligned}$$

The factor  $\binom{j}{s}$  in the first equation is because as long as we choose  $s$  positions for children coming from the deterministic component of offspring distribution, the order of all  $j$  children are fixed since the order among  $s$  children and the order among the other  $j - s$  are both fixed.  $\square$

We remark that if  $\mathcal{T}$  has deterministic  $d$ -ary branching (every node has exactly  $d$  children with probability one), then for all  $j$ , the number of subtrees containing the root and having precisely  $j$  nodes is exactly  $\binom{dj}{j-1}/j$  (Thm 5.3.10 in [19]) which is bounded above by  $\left(\frac{edj}{j-1}\right)^{j-1}/j \leq (ed)^j$ . Thus, Lemma 6 shows that when factorial moments grow only exponentially quickly, the values  $\mu_j$  behave roughly as in the case of deterministic branching. We also note that when  $\mathcal{T}$  has  $\text{Poisson}(c)$  branching distribution, Lemma 6 and the argument of Lemma 7 together yield the exact formula  $\mu_j = \frac{c^{j-1}}{j} \binom{2j-2}{j-1}$ .

*Proof of Proposition 3.* For any  $v \in V(H_{n,k,p})$ , the random variable  $|B_{j,v}|$  is stochastically dominated by  $\mu_j(B)$ , where  $B \stackrel{d}{=} \text{Bin}(n, c/n) + 2k$ . By Lemma 7, we have that  $q_j(B) \leq (c+2k)^j$  for all  $j$ . It then follows from Lemma 6 that  $q_j \leq (4(c+2k))^{j-1}$  for all  $j$ , proving the proposition.  $\square$

#### 4. BOUNDING THE EXPANSION OF CONNECTED SUBGRAPHS OF $H_{n,k,p}$

Recall from the preceding section that  $B_j$  is the collection of connected subsets  $S$  of  $V(H_{n,k,p})$  with  $|S| = j$ . We will show that with high probability, for all  $j \geq \log n$ , all elements  $S$  of  $B_j$  have conductance uniformly bounded away from zero. Many of our proofs are easier when  $c$  is large, and we treat this case first.

In the course of the proofs we will make regular use of the standard Chernoff bounds (see, e.g., [8] Theorem 2.1) which we summarize here:

**Theorem 8.** *If  $X \stackrel{d}{=} \text{Bin}(m, q)$  then*

$$\text{for all } 0 < x < 1, \mathbf{P}(X \leq (1-x)mq) \leq \exp(-mq\phi(-x)) \leq \exp(-mqx^2/2),$$

$$\text{and for all } x > 0, \mathbf{P}(X \geq (1+x)mq) \leq \exp(-mq\phi(x)) \leq \exp(-mqx^2/2(1+x)),$$

$$\text{where } \phi(x) = (1+x)\log(1+x) - x.$$

We will use the coarser bounds most of the time. The finer bound will only be used twice, once in the proof of Lemma 13 and once in the proof of Theorem 1.

We will also have use of the FKG inequality ([8], Theorem 2.12), which we now recall. Let  $\Gamma = [n] = \{1, 2, \dots, n\}$ . Given  $0 \leq p_1, \dots, p_n \leq 1$ ,  $\Gamma_{p_1, \dots, p_n} \subset [n]$  is obtained by including



element  $i$  with probability  $p_i$  independently for all  $i$ . We say a function  $f : 2^\Gamma \rightarrow \mathbb{R}$  is *increasing* if  $f(A) \leq f(B)$  for  $A \subset B$ ,  $f$  *decreasing* if  $f(A) \geq f(B)$  for  $A \subset B$ .

**Theorem 9** (FKG inequality). *If the random variables  $X_1$  and  $X_2$  are two increasing or two decreasing functions of  $\Gamma_{p_1, \dots, p_n}$ , then*

$$\mathbb{E}(X_1 X_2) \geq \mathbb{E}(X_1) \mathbb{E}(X_2).$$

*In particular, if  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$  are two increasing or two decreasing families of subsets of  $\Gamma$ , then*

$$\mathbb{P}(\Gamma_{p_1, \dots, p_n} \in \mathcal{Q}_1 \cap \mathcal{Q}_2) \geq \mathbb{P}(\Gamma_{p_1, \dots, p_n} \in \mathcal{Q}_1) \mathbb{P}(\Gamma_{p_1, \dots, p_n} \in \mathcal{Q}_2).$$

We will in fact use the following, easy consequence of the FKG inequality.

**Corollary 10.** *If  $C$  is an increasing event,  $A$  is a decreasing event, then*

$$\mathbf{P}(C \mid A) \leq \mathbf{P}(C).$$

We begin by bounding the edge expansion of all but the very large connected sets, in the case that  $c$  is large.

**Lemma 11.** *Fix  $c$  sufficiently large that  $c/720 - \log(4(c+2k)) > 5$ . Then for all  $n$ ,*

$$\mathbf{P}\left(\exists S \in \bigcup_{\log n \leq j \leq 9n/10} B_j, e(S, S^c) \leq c|S|/12\right) \leq \frac{1}{n^3}.$$

*Proof.* Fix  $j \in [\log n, 9n/10]$ , and  $S \subset [n]$  with  $|S| = j$ . Note that  $E(S, S^c)$  is independent of  $H[S]$ , so given that  $S \in B_j$ ,  $e(S, S^c)$  stochastically dominates a  $\text{Bin}(j(n-j), p)$  random variable. Since  $j(n-j) \geq jn/10$ , it follows that under this conditioning  $e(S, S^c)$  also stochastically dominates  $X$ , a  $\text{Bin}(nj/10, p)$  random variable. By a union bound it follows that

$$\begin{aligned} & \mathbf{P}(\exists S, |S| = j, H[S] \text{ connected}, e(S, S^c) \leq cj/12) \\ & \leq \sum_{S, |S|=j} \mathbf{P}(e(S, S^c) \leq cj/12 \mid H[S] \text{ connected}) \cdot \mathbf{P}(H[S] \text{ connected}) \\ & \leq \mathbf{P}(X \leq cj/12) \cdot \mathbf{E}|B_j| \\ & \leq e^{-cj/720} \cdot n(4(c+2k))^j, \end{aligned}$$

the last line by a Chernoff bound and by Proposition 3. (This is a typical example of our use of Proposition 3 in the remainder of the paper.)

By our assumption that  $c/720 - \log(4(c+2k)) > 5$  and since  $j \geq \log n$ , we obtain that this probability is at most

$$\exp\left(\log n + j\left(\log(4(c+2k)) - \frac{c}{720}\right)\right) \leq \frac{1}{n^4}.$$

The result follows by a union bound over  $j \in [\log n, 9n/10]$ .  $\square$

The next lemma provides a lower bound on the edge expansion of very large sets, again in the case that  $c$  is sufficiently large.

**Lemma 12.** *If  $c > 40$  then for all  $n$  sufficiently large,*

$$\mathbf{P}(\exists S \subset [n] : |S| > 9n/10, e(S) \leq |E(H)|) \leq (2/e)^n.$$



*Proof.* In this proof write  $E = E(H)$ . Since, for any set  $S \subset [n]$ ,  $e(S) + e(S^c) = 2|E|$ , it suffices to prove that

$$\mathbf{P}(\exists S \subset [n] : |S| < n/10, e(S) \geq |E|) \leq (2/e)^n.$$

Fix any set  $S$  with  $|S| < n/10$ . Write  $e^*(S) = \sum_{v \in S} |\{e \ni v : e \notin E(R_{n,k})\}|$  for the total degree incident to  $S$  not including edges of the ring  $R_{n,k}$ , and similarly let  $E^* = E \setminus E(R_{n,k})$ . Since  $|S| < n/10 < n/2$ , in order to have  $e(S) \geq |E|$  we must in fact have  $e^*(S) \geq |E^*|$ . Also,  $e^*(S)$  is stochastically dominated by  $\text{Bin}(n^2/10, p)$ , and  $|E^*| \stackrel{d}{=} \text{Bin}(n(n-1-2k)/2, p)$ . When  $n$  is large enough that  $n-1-2k > 4n/5$ , we thus have

$$\begin{aligned} \mathbf{P}(e(S) \geq |E(H)|) &\leq \mathbf{P}(|E^*| \leq cn/5) + \mathbf{P}(e^*(S) > cn/5) \\ &\leq \mathbf{P}(|E^*| \leq c(n-1-2k)/4) + \mathbf{P}(e^*(S) > cn/5) \\ &< \exp(-c(n-1-2k)/16) + \exp(-cn/40), \end{aligned}$$

by a Chernoff bound. For  $n$  sufficiently large the last line is at most  $2e^{-cn/40} < 2e^{-n}$ , and the result follows by a union bound over all  $S$  with  $|S| \leq n/10$  (there are less than  $2^{n-1}$  such sets).  $\square$

A similar but slightly more involved argument yields the following result, which will be useful for dealing with smaller values of  $c$ .

**Lemma 13.** *For any  $c > 0$  there is  $\beta = \beta(c) > 0$  such that for all  $n$  sufficiently large*

$$\mathbf{P}(\exists S \subset [n] : |S| > (1-\beta)n, e(S) \leq |E(H)|) \leq (1-\beta)^n.$$

*Proof.* As in the proof of Lemma 12, it suffices to prove that for some  $\beta > 0$ ,

$$\mathbf{P}(\exists S \subset [n] : |S| < \beta n, e(S) > |E|) \leq (1-\beta)^n.$$

Furthermore, since  $\mathbf{P}(\exists S \subset [n] : |S| < \beta n, e(S) > |E|)$  decreases as  $\beta$  decreases, it suffices to find  $\beta > 0$  and  $\epsilon > 0$  such that for  $n$  sufficiently large,

$$\mathbf{P}(\exists S \subset [n] : |S| < \beta n, e(S) > |E|) = O(e^{-\epsilon n}).$$

We fix  $0 < \beta < 1/(3e)$  small enough that  $1/(2\beta) - 8k/c > 1 + 1/(3\beta)$ . Additionally, recalling the function  $\phi(x) = (1+x)\log(1+x) - x$  from Theorem 8, we choose  $\beta$  small enough that  $\phi(1/(3\beta)) > \log(1/(3\beta))/(6\beta)$ . Finally, we assume  $\beta < c/36$ .

For any  $S \subset [n]$  with  $|S| < \beta n$  and with  $e(S) > |E|$ , defining  $e^*(S)$  and  $E^*$  as in the proof of Lemma 12, we then have  $e^*(S) \geq |E| - 2k|S| \geq |E| - 2k\beta n$ .

$$\begin{aligned} &\mathbf{P}(\exists S \subset [n] : |S| < \beta n, e(S) > |E|) \\ &\leq \mathbf{P}(\exists S \subset [n] : |S| < \beta n, e^*(S) \geq |E| - 2k\beta n) \\ &\leq \mathbf{P}\left(|E| \leq \binom{n}{2}p - 2k\beta n\right) + \mathbf{P}\left(\exists S \subset [n] : |S| < \beta n, e^*(S) > \binom{n}{2}p - 4k\beta n\right). \end{aligned} \quad (3)$$

Since  $|E|$  stochastically dominates  $\text{Bin}(\binom{n}{2}, p)$ , by a Chernoff bound we have

$$\begin{aligned} \mathbf{P}\left(|E| \leq \binom{n}{2}p - 2k\beta n\right) &\leq \exp\left(-\frac{1}{2}\binom{n}{2}p\left(\frac{2k\beta n}{\binom{n}{2}p}\right)^2\right) \\ &< \exp\left(-\frac{4k^2\beta^2}{c} \cdot n\right), \end{aligned} \quad (4)$$

which handles the first summand in (3). For the second summand, let  $X$  be  $\text{Binomial}(\beta n(n-1), p)$  distributed, and note that for all  $S \subset [n]$  with  $|S| \leq \beta n$ ,  $e^*(S)$  is stochastically dominated

by  $X$ . Also, for  $\beta < 1/3$  there are less than  $2\binom{n}{\lfloor \beta n \rfloor}$  subsets of  $[n]$  of size less than  $\beta n$ , and follows by a union bound that

$$\mathbf{P}\left(\exists S \subset [n] : |S| < \beta n, e^*(S) > \binom{n}{2}p - 4k\beta n\right) \leq 2\binom{n}{\lfloor \beta n \rfloor} \mathbf{P}\left(X > \binom{n}{2}p - 4k\beta n\right). \quad (5)$$

Since  $1/(p(n-1)) = n/(c(n-1)) \leq 2/c$  for all  $n \geq 2$ , and by our assumption that  $1/(2\beta) - 8k/c > 1 + 1/(3\beta)$ , we have

$$\binom{n}{2}p - 4k\beta n = \beta n(n-1)p \left( \frac{1}{2\beta} - \frac{4k}{p(n-1)} \right) > \beta n(n-1)p \left( 1 + \frac{1}{3\beta} \right).$$

By the sharper of the Chernoff upper bounds in Theorem 8 and by our assumption that  $\phi(1/(3\beta)) > \log(1/(3\beta))/(6\beta)$ , it follows that

$$\begin{aligned} \mathbf{P}\left(X > \binom{n}{2}p - 4k\beta n\right) &\leq \exp(-\beta n(n-1)p \cdot \phi(1/(3\beta))) \\ &\leq \exp\left(-\beta n(n-1)p \frac{\log(1/(3\beta))}{6\beta}\right) \\ &< \exp\left(-\frac{c \log(1/(3\beta))}{12}n\right) \end{aligned}$$

for  $n$  sufficiently large.

Combined with (5) this yields

$$\begin{aligned} &\mathbf{P}\left(\exists S \subset [n] : |S| < \beta n, e^*(S) > \binom{n}{2}p - 4k\beta n\right) \\ &< 2\binom{n}{\lfloor \beta n \rfloor} \exp\left(-\frac{c \log(1/(3\beta))}{12}n\right) \\ &\leq 2\left(\frac{e}{\beta}\right)^{\beta n} \exp\left(-\frac{c \log(1/(3\beta))}{12}n\right) \\ &= 2\exp(n(\beta + \beta \log(1/\beta) - (c/12)\log(1/(3\beta)))) \\ &< 2\exp(-(c/36)n), \end{aligned}$$

where in the last inequality we used that  $\beta + \beta \log(1/\beta) < 2\beta \log(1/\beta) < (c/18)\log(1/\beta)$  and that  $\log(1/(3\beta)) > 1$ . Together with (3) and (4) we obtain

$$\mathbf{P}(\exists S \subset [n] : |S| < \beta n, e(S) > |E|) \leq \exp\left(-\frac{4k^2\beta^2}{c} \cdot n\right) + 2\exp(-(c/36)n),$$

which completes the proof.  $\square$

In order to use Lemma 11 to bound the conductance of connected subsets  $S$  of  $V(H_{n,k,p})$  of size at most  $9n/10$ , we need to know that for such subsets we have  $e(S) = O(|S|)$  with high probability. Such a bound is provided by the following lemma.

For given  $k$ , let  $x = x_k$  be the positive solution of equation  $x/720 - \log(4(x+2k)) = 5$ , and let  $M = M(c, k) = k + 1 + 10\max(x_k, c)$ . We remark that  $x_k \geq 40$  for all  $k \geq 1$ .

**Lemma 14.** *For all  $c > 0$  and for all  $n$ ,*

$$\mathbf{P}\left(\exists S \in \bigcup_{1 \leq j \leq n} B_j, e(S, S) > M \cdot \max(|S|, \log n)\right) \leq \frac{1}{n^3}.$$

*Proof.* First note that for *fixed*  $M$ , the event whose probability we aim to bound is increasing, so increasing  $p$  only increases its probability of occurrence. Since  $M(c, k)$  is constant for  $c \leq x_k$ , it thus suffices to prove the bound for  $c \geq x_k$ , and we now assume that  $c \geq x_k$ . Note that in this case  $c/720 - \log(4(c + 2k)) \geq 5$ . For all  $n$ , and any  $j \in [n]$ , we have

$$\begin{aligned} & \mathbf{P}(\exists S \in B_j, e(S, S) > (k + 1 + 10c) \max(j, \log n)) \\ & \leq \sum_{S \subset [n], |S|=j} \mathbf{P}(e(S, S) > (k + 1 + 10c) \max(j, \log n), S \in B_j). \end{aligned} \quad (6)$$

Write  $\mathbf{T}_S$  for the set of all possible trees on vertex set  $S$  – so  $|\mathbf{T}_S| = |S|^{|S|-2}$  – and list the elements of  $\mathbf{T}_S$  as  $t_1, \dots, t_r$ . For  $i \in [r]$ , let  $F_i$  be the event that  $t_i$  is a subgraph of  $H$ , and let  $E_i = F_i \setminus \bigcup_{j < i} F_j$  be the event that  $t_i$  is a subgraph of  $H$  but none of  $t_1, \dots, t_{i-1}$  are subgraphs of  $H$ . The events  $E_i$  partition the event that  $S \in B_j$ , so

$$\begin{aligned} & \mathbf{P}(e(S, S) > (k + 1 + 10c) \max(j, \log n) \mid S \in B_j) \\ & \leq \max_{i \in [r]} \mathbf{P}(e(S, S) > (k + 1 + 10c) \max(j, \log n) \mid E_i) \end{aligned} \quad (7)$$

For fixed  $i \in [r]$ , write  $\mathbf{P}_i(\cdot)$  for the conditional probability measure  $\mathbf{P}(\cdot \mid F_i)$ , and write  $\Gamma^{(i)} = \{uv : u, v \in S\} \setminus E(t_i)$ . Under  $\mathbf{P}_i$ , the set  $E(S, S) \setminus E(t_i)$  is distributed as a Binomial( $p$ ) random subset of  $\Gamma^{(i)}$  since, after conditioning that  $t_i$  is a subgraph of  $H$ , those edges not in  $t_i$  still appear independently.

Write  $C$  for the event that  $|E(S, S) \setminus E(t_i)| > |(k + 1 + 10c) \max(j, \log n)| - (j - 1)$ . Then

$$\mathbf{P}(e(S, S) > (k + 1 + 10c) \max(j, \log n) \mid E_i) = \mathbf{P}_i\left(C \mid \bigcap_{j < i} F_j^c\right). \quad (8)$$

Since  $C$  is increasing and  $\bigcap_{j < i} F_j^c$  is decreasing, it follows from Corollary 10 that

$$(8) \leq \mathbf{P}_i(C) = \mathbf{P}(|E(S, S) \setminus E(t_i)| > (k + 1 + 10c) \max(j, \log n) - (j - 1)),$$

the last equality holding since  $E(S, S) \setminus E(t_i)$  is disjoint from  $E(t_i)$ , so independent of  $F_i$ . Now,  $|E(S, S) \setminus E(t_i)|$  is stochastically dominated by  $kj + \text{Bin}(\max(j, \log n) \cdot n/2, p)$ . Letting  $X$  have distribution  $\text{Bin}(n \max(j, \log n)/2, p)$ , respectively, for all  $i \in [r]$  we thus have

$$\begin{aligned} & \mathbf{P}(e(S, S) > (k + 1 + 10c) \max(j, \log n) \mid E_i) \\ & \leq \mathbf{P}(kj + (j - 1) + X > (k + 1 + 10c) \max(j, \log n)) \\ & \leq \mathbf{P}(X > 10c \max(j, \log n)) \\ & \leq e^{-9c \max(j, \log n)/2}, \end{aligned}$$

the last inequality holding by a Chernoff bound. It then follows from (6) and (7) that

$$\begin{aligned} & \mathbf{P}(\exists S \in B_j, e(S, S) > (k + 1 + 10c) \max(j, \log n)) \\ & \leq \sum_{S \subset [n], |S|=j} e^{-9c \max(j, \log n)/2} \mathbf{P}(S \in B_j) \\ & = e^{-9c \max(j, \log n)/2} \mathbf{E}|B_j| \\ & \leq e^{-9c \max(j, \log n)/2} n(4(c + 2k))^j \end{aligned}$$

the last inequality by Proposition 3. By assumption,  $c$  is large enough that  $c/720 - \log(4(c + 2k)) > 5$ , and it follows that

$$\begin{aligned} & \mathbf{P}(\exists S \in B_j, e(S, S) > (k + 1 + 10c) \max(j, \log n)) \\ & \leq e^{\log n + \log(4(c+2k)) \max(j, \log n) - 9c \max(j, \log n)/2} \\ & < n^{-4}. \end{aligned}$$

A union bound over  $j \in [n]$  completes the proof.  $\square$

## 5. PROOF OF THEOREM 1

As noted in the introduction, the case when  $c$  is large is more straightforward, and we handle it first.

*Proof of Theorem 1 assuming  $c > x_k$ .* For  $x \geq 0$  write

$$\Phi_0(x) = \min \left\{ \frac{e(S, S^c)}{e(S)} : S \text{ connected}, |S| \leq \frac{9n}{10}, xn(c/2 + k)/2 \leq e(S) \leq 4xn(c/2 + k) \right\}.$$

Let  $A$  be the event that for all  $S$  with  $|S| > 9n/10$  we have  $e(S) > |E(H)|$ . If  $A$  occurs then for all  $0 \leq x \leq 1/2$ , for all  $S$  with  $e(S) \leq 2x|E(H)|$  we have  $e(S) \leq |E(H)|$  so  $|S| \leq 9n/10$ . Also, let  $A'$  be the event that  $n(c/2 + k)/2 \leq |E(H)| \leq 2n(c/2 + k)$ . If  $A'$  occurs then for all  $x \geq 0$  all  $S$  with  $x|E(H)| \leq e(S) \leq 2x|E(H)|$ , we have  $xn(c/2 + k)/2 \leq e(S) \leq 4xn(c/2 + k)$ . It follows that on  $A \cap A'$ , for all  $0 \leq x \leq 1/2$  we have  $\Phi_0(x) \leq \Phi(x)$ .

We have  $|E(H)| \stackrel{d}{=} nk + \text{Bin}(n(n - 2k - 1)/2, p)$ , so by a Chernoff bound, for all  $n$  large enough,  $\mathbf{P}(A') \geq 1 - n^{-3}$ . Also, by Lemma 12,  $\mathbf{P}(A) \geq 1 - n^{-3}$  for all  $n$  sufficiently large. It follows that with probability at least  $1 - 2n^{-3}$ , for all  $0 \leq x \leq 1/2$  we have  $\Phi_0(x) \leq \Phi(x)$ , and thus with probability at least  $1 - 2n^{-3}$ ,

$$\sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi^{-2}(2^{-i}) \leq \sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi_0^{-2}(2^{-i}).$$

We thus focus on bounding the latter quantity. Note that the sets considered when bounding  $\Phi(2^{-i})$  decrease (in terms of  $e(S)$ ) as  $i$  increases. Also, recall the definitions of  $x_k$  and  $M = M(c, k)$  from just before the statement of Lemma 14.

First suppose  $i \geq \lfloor \log_2(n(c/2 + k)/(8M \log n)) \rfloor$  and write

$$i = \lfloor \log_2(n(c/2 + k)/(8M \log n)) \rfloor + j.$$

For all  $S$  considered when bounding  $\Phi_0(2^{-i})$  we have  $e(S) \leq 2^{-j+6} M \log n$  and  $e(S, S^c) \geq 2$ , so

$$\begin{aligned} & \sum_{j=1}^{\lceil \log_2 |E| \rceil - \lfloor \log_2(n(c/2 + k)/(8M \log n)) \rfloor} \Phi_0^{-2}(2^{-\lfloor \log_2(n(c/2 + k)/(8M \log n)) \rfloor - j}) \\ & \leq 4^5 M^2 \log^2 n \sum_{j=0}^{\infty} 4^{-j} = \frac{4^6 M^2}{3} \log^2 n. \end{aligned}$$

Next suppose that  $i \leq \log_2(n(c/2 + k)/(8M \log n))$ . In this case we have  $e(S) \geq 4M \log n$ . By Lemma 14, with probability at least  $1 - n^{-3}$ , for all connected sets  $S$  with  $|S| \leq \log n$  we have  $e(S, S) \leq M \log n$ . Since  $e(S) \geq 4M \log n$  this implies that  $e(S, S^c) = e(S) - 2e(S, S) \geq$

$e(S) - 2M \log n \geq e(S)/2$ , so  $\Phi(S) \geq 1/2$ . Also, by Lemmas 11 and 14, with probability at least  $1 - 2n^{-3}$ , for all connected sets  $S$  with  $|S| \geq \log n$ , we have

$$\frac{e(S, S^c)}{e(S)} = \frac{e(S, S^c)}{e(S, S^c) + 2e(S, S)} \geq \frac{c|S|}{12} \frac{1}{c|S|/12 + 2M|S|} \geq \frac{c}{36M},$$

so  $\Phi(S) \geq c/(36M)$ . Since  $1/2 > c/(36M)$  it follows that with probability at least  $1 - 3n^{-3}$ , for all  $i \leq \log_2(n(c/2 + k)/(8M \log n))$  we have  $\Phi_0(2^{-i}) \geq \frac{c}{36M}$ , and in this case

$$\sum_{i=1}^{\lfloor \log_2(n(c/2+k)/(8M \log n)) \rfloor} \Phi_0^{-2}(2^{-i}) \leq \frac{6^4 M^2}{c^2} \log_2 n.$$

Combining these bounds, we see that with probability at least  $1 - 3n^{-3}$ ,

$$\sum_{i=1}^{\lfloor \log_2 |E| \rfloor} \Phi_0^{-2}(2^{-i}) \leq \frac{4^6 M^2}{3} \log^2 n + \frac{6^4 M^2}{c^2} \log_2 n,$$

so with probability at least  $1 - 5n^{-3}$ ,  $\sum_{i=1}^{\lfloor \log_2 |E| \rfloor} \Phi^{-2}(2^{-i})$  is at most the same quantity. By Theorem 2 it follows that with probability at least  $1 - 5n^{-3}$ ,

$$\tau_{\text{MIX}}(G) \leq C \left( \frac{4^6 M^2 \log^2 n}{3} + \frac{6^4 M^2}{c^2} \log_2 n \right).$$

This completes the proof in the case  $c > x_k$ .  $\square$

For the remainder of the paper, fix  $0 < c < x_k$ , and let  $R = \lceil \max(k, 2x_1/c) \rceil$ . Also, recall the constant  $\beta = \beta(c)$  from Lemma 13. The remaining case of Theorem 1 follows straightforwardly from the following lemma.

**Lemma 15.** *There is  $\alpha = \alpha(c) > 0$  such that for all  $n$  sufficiently large,*

$$\mathbf{P} \left( \exists S \in \bigcup_{R \log n \leq j \leq (1-\beta)n} B_j(H), e_H(S, S^c) \leq \alpha|S| \right) \leq \frac{3R^3}{n^3}.$$

We provide the proof of Lemma 15 at the end of the paper.

*Proof of Theorem 1 assuming  $c \leq x_k$ .* For  $x \geq 0$  write

$$\Phi_0(x) = \min \left\{ \frac{e(S, S^c)}{e(S)} : S \text{ connected}, |S| \leq (1-\beta)n, xn(c/2 + k)/2 \leq e(S) \leq 4xn(c/2 + k) \right\}.$$

As in the case  $c > x_k$ , by a Chernoff bound and by Lemma 13, for all  $n$  sufficiently large, with probability at least  $1 - 2n^{-3}$  we have

$$\sum_{i=1}^{\lfloor \log_2 |E| \rfloor} \Phi^{-2}(2^{-i}) \leq \sum_{i=1}^{\lfloor \log_2 |E| \rfloor} \Phi_0^{-2}(2^{-i}),$$

and the remainder of the proof is just as in the case  $c > x_k$ , but using Lemma 15 in place of Lemma 11.  $\square$

It thus remains to prove Lemma 15; before doing so, we briefly describe our approach. We shall divide vertices of  $H$  into groups of size  $R$ , each containing  $R$  consecutive vertices. We view each group as a new single vertex; two new vertices are connected if there is an edge connecting their constituent sets. This yields an auxiliary graph  $H'$ , whose distribution is that of an  $(n', 1, p')$  Newman-Watts small world, for suitable  $n'$  and  $p'$ . We will shortly see that  $p' = c'/n'$  for some  $c' > x_1$ , so all “large- $c$ ” results can be applied to  $H'$ .

To translate edge expansion results from  $H'$  into corresponding results for  $H$ , we proceed as follows. Given a set  $S$  of vertices of  $H$ , we consider the *blow-up*  $S^+$  of  $S$ , which is the collection of all vertices of  $H$  belonging to the same group as some element of  $S$ . The idea is that in most cases, the event that  $e(S, S^c)$  is small relative to  $|S|$  should be nearly identical to the event that  $e(S^+, (S^+)^c)$  is small relative to  $|S^+|$ . If this were always true, Lemma 11 would then yield bounds for the number of edges leaving  $S^+$ , which would in turn yield strong bounds on the probability that  $e(S^+, (S^+)^c) < \epsilon |S^+|$ , where  $\epsilon > 0$  will be a function of  $R$ .

The above line of argument relies upon the intuition that the size of the blow-up  $S^+$  should be essentially a constant factor greater than that of  $S$ . Since the ratio  $|S^+|/|S|$  is in fact a random quantity, to make the above argument work, we end up needing to additionally show that  $e(S^+, (S^+)^c)$  is very unlikely to be large if  $e(S, S^c)/|S|$  is extremely small. In order to quantify the notion of “extremely small”, we are forced to introduce a third parameter  $\delta > 0$  with  $\delta$  much smaller than  $\epsilon$ . We now turn to the details.

*Proof of Lemma 15.* We assume for simplicity that  $R$  divides  $n$  – the general case is practically identical – and write  $n' = n/R$ . For  $i \in [n']$  let  $w_i = \{(i-1)R + j, 1 \leq j \leq R\}$ . We form an auxiliary graph  $H' = (V', E')$  with  $V' = \{w_i, i \in [n']\}$  by adding an edge between  $w_i$  and  $w_j$  if there is some edge from an element of  $w_i$  to an element of  $w_j$  in  $H$ . It is easily verified (using the fact that  $R > k$ ) that  $H'$  is an  $(n', 1, p')$  Newman–Watts small world, with  $p' = \mathbf{P}(\text{Bin}(R^2, p) > 0) > Rc/2n' = c'/n'$  (where  $c' = Rc/2$ ) for all  $n$  sufficiently large. Note that since  $c' = Rc/2 > x_1$ , it follows that we may apply Lemma 11 to  $H'$  – this is the only way we will use this bound on  $R$ .

Given  $S \subset [n]$ , write  $I' = \{i \in n' : w_i \cap S \neq \emptyset\}$ , let  $S' = \{w_i, i \in I'\}$ , and write  $S^+ = \bigcup_{i \in S'} w_i$ . Now write  $j = |S|$ . As in the proof of Lemma 14, we partition the event that  $S \in B_j(H)$  into  $E_1(S), \dots, E_r(S)$  according to the first spanning tree appearing in  $S$ . Fix  $\epsilon = c/(12R(2Rc+1))$  and let  $\delta > 0$  be small enough that  $\epsilon c \geq 2k\delta$  and that  $\epsilon Rc(\log(\frac{\epsilon}{\delta}) - 1) \geq 5 + \log(4(c+2k))$ . For all  $1 \leq i \leq r$  we then have

$$\begin{aligned} & \mathbf{P}(e_H(S^+ \setminus S, (S^+)^c) > 2\epsilon Rcj \mid e_H(S, S^c) \leq \delta j, E_i(S)) \\ & \leq \mathbf{P}(e_H(S^+ \setminus S, (S^+)^c) > (\epsilon c + 2k\delta)Rj \mid e_H(S, S^c) \leq \delta j, E_i(S)) \\ & \leq \mathbf{P}(e_H(S^+ \setminus S, (S^+)^c) > (\epsilon c + 2k\delta)Rj \mid e_H(S, S^c) \leq \delta j, t_i \subset H). \end{aligned}$$

The first inequality is true since we pick  $\delta$  such that  $\epsilon c \geq 2k\delta$  and the second inequality holds by Corollary 10. Now write  $g(S) = |\{i \in S' : |S \cap w_i| < |w_i|\}|$ , so  $g(S)$  is the number of sets  $w_i$  that intersect  $S$  but are not covered by  $S$ . It is easily checked that  $e_H(S, S^c) \geq g(S)$ ; the extremal case is that for each  $i \in S'$ ,  $e_H(S \cap w_i, w_i \setminus S) = 1$ , while for  $i, j \in S'$  with  $i \neq j$ ,  $e_H(w_i, w_j) = 0$ .

Next, suppose that  $S \subset [n]$  satisfies  $e_H(S, S^c) \leq \delta j$ . Then we must have  $g(S) \leq \delta j$ , and it follows that  $|S^+ \setminus S| \leq R\delta j$ . Under this conditioning,  $e_H(S^+ \setminus S, (S^+)^c)$  is stochastically dominated by  $2kR\delta j + \text{Bin}(R\delta jn, p)$ , and is independent of the event that  $t_i \subset H$  since they are determined by disjoint sets of edges, so by the finer of the Chernoff upper bounds,

$$\mathbf{P}(e_H(S^+ \setminus S, (S^+)^c) > (\epsilon c + 2k\delta)Rj \mid e_H(S, S^c) \leq \delta j, t_i \subset H) \leq \exp\left(-\epsilon Rcj\left(\log\left(\frac{\epsilon}{\delta}\right) - 1\right)\right).$$

It follows that for  $j \geq R \log n$ , writing

$$S_1 = \{S \subset [n], S \in B_j(H), e_H(S, S^c) \leq \delta j, e_H(S^+ \setminus S, (S^+)^c) > 2\epsilon Rcj\},$$

we have

$$\begin{aligned}
\mathbf{E}|S_1| &\leq \sum_{S \subset [n], |S|=j} \sum_{1 \leq i \leq r} \mathbf{P}(E_i(S), e_H(S, S^c) \leq \delta|S|, e_H(S^+ \setminus S, (S^+)^c) > 2\epsilon Rc|S|) \\
&\leq \exp\left(-\epsilon Rcj \left(\log\left(\frac{\epsilon}{\delta}\right) - 1\right)\right) \mathbf{E}|\{S \subset [n], |S|=j, H[S] \text{ connected}\}| \\
&\leq \exp\left(-\epsilon Rcj \left(\log\left(\frac{\epsilon}{\delta}\right) - 1\right)\right) \cdot n \cdot (4(c+2k))^j \\
&\leq n^{-4}
\end{aligned}$$

the second-to-last inequality by Proposition 3, and the last since we chose  $\delta$  such that

$$\epsilon Rc \left(\log\left(\frac{\epsilon}{\delta}\right) - 1\right) \geq 5 + \log(4(c+2k)).$$

and since  $j \geq R \log n \geq \log n$ . It follows by a union bound over  $R \log n \leq j \leq 9n/10$  and Markov's inequality that

$$\mathbf{P}\left(\exists S \in \bigcup_{R \log n \leq j \leq 9n/10} B_j(H), e_H(S, S^c) \leq \delta|S|, e_H(S^+ \setminus S, (S^+)^c) > 2\epsilon Rc|S|\right) \leq \frac{1}{n^3}. \quad (9)$$

Next, write

$$\mathcal{S}_2 = \left\{ S \in \bigcup_{R \log n \leq j \leq 9n/(10R)} B_j(H), e_H(S, S^c) \leq \epsilon|S|, e_H(S^+ \setminus S, (S^+)^c) \leq 2\epsilon Rc|S| \right\}.$$

For any  $S \in \mathcal{S}_2$ , we have

$$\begin{aligned}
e_H(S^+, (S^+)^c) &= e_H(S, (S^+)^c) + e_H(S^+ \setminus S, (S^+)^c) \\
&\leq e_H(S, S^c) + e_H(S^+ \setminus S, (S^+)^c) \\
&\leq \epsilon(2Rc + 1)|S| \\
&\leq \epsilon(2Rc + 1)R|S'| \\
&= c'|S'|/12.
\end{aligned}$$

the last equality by our choice of  $\epsilon$ . It follows that  $e_{H'}(S', (S')^c) \leq c'|S'|/12$ . Furthermore, since  $|S| \geq R \log n$  we have  $|S'| \geq \log n \geq \log n'$ , and since  $|S| \leq 9n/(10R)$  we have  $|S'| \leq 9n/(10R) = 9n'/10$ . It follows by Lemma 11 that

$$\begin{aligned}
&\mathbf{P}\left(\exists S \in \bigcup_{R \log n \leq j \leq 9n/(10R)} B_j(H), e_H(S, S^c) \leq \epsilon|S|, e_H(S^+ \setminus S, (S^+)^c) \leq 2\epsilon Rc|S|\right) \\
&\leq \mathbf{P}\left(\exists S' \in \bigcup_{\log n' \leq j \leq 9n'/10} B_j(H'), e_{H'}(S', (S')^c) \leq c'|S'|/12\right) \leq \frac{1}{(n')^3} \quad (10)
\end{aligned}$$

Next, for any  $m \geq 1$ , if  $e_H(S, S^c) \leq m$  then, viewed as a subset of a cycle of length  $n$ ,  $S$  must have at most  $m$  connected components. The number of subsets of an  $n$ -cycle with at most  $m$  connected components is  $2n \binom{n+2m-1}{2m-1}$ . (This is a straightforward combinatorial exercise but may be seen as follows: the factor  $n$  chooses a starting point on the cycle, the factor  $\binom{n+2m-1}{2m-1}$  chooses the points on the cycle at which membership in  $S$  alternates, and the factor 2 accounts for whether or not the starting point belongs to  $S$ .)



It follows that for any  $\gamma > 0$ , the number of subsets of an  $n$ -cycle with at most  $\gamma n$  connected components is at most

$$\begin{aligned} 2n \binom{n + \lfloor 2\gamma n \rfloor}{\lfloor 2\gamma n \rfloor} &\leq 2n \left( \frac{e(1+2\gamma)n}{2\gamma n} \right)^{2\gamma n} \\ &= \exp \left( \log(2n) + n \cdot 2\gamma \left( 1 + \log \left( 1 + \frac{1}{2\gamma} \right) \right) \right) \end{aligned} \quad (11)$$

Since  $x(1 + \log(1/(2x))) \rightarrow 0$  as  $x \downarrow 0$ , we may choose  $0 < \gamma < 9\beta c/(20R)$  small enough that (11) is at most  $\exp(n \cdot 9\beta c/(160R))$  for all  $n$  sufficiently large.

Finally, for a fixed set  $S$  with  $9n/(10R) \leq |S| \leq (1 - \beta)n$ , the probability that  $e_H(S, S^c) \leq \gamma n$  is bounded above by  $\mathbf{P} \left( \text{Bin}(\frac{9n}{10R} \cdot \beta n, p) \leq \gamma n \right)$  since  $e_H(S, S^c)$  stochastically dominates  $\text{Bin}(|S|(n - |S|), p)$  and  $|S| \geq \frac{9n}{10R}$ ,  $n - |S| \geq \beta n$ . Since  $\frac{10R\gamma}{9\beta c} \leq \frac{1}{2}$ , by a Chernoff bound we have

$$\mathbf{P}(e_H(S, S^c) \leq \gamma n) \leq \mathbf{P} \left( \text{Bin}(\frac{9n}{10R} \cdot \beta n, p) \leq \gamma n \right) \leq \exp \left( -\frac{9\beta c}{80R} n \right). \quad (12)$$

Since *all* sets  $S \subset [n]$  with at least  $\gamma n$  components (still viewed as subsets of the  $n$ -cycle) have  $e(S, S^c) \geq \gamma n$ , it follows by (11), (12), and a union bound over sets with at most  $\gamma n$  components that for all  $n$  sufficiently large,

$$\begin{aligned} \mathbf{P} \left( \exists S \in \bigcup_{9n/(10R) \leq j \leq (1-\beta)n} B_j(H), e_H(S, S^c) \leq \gamma n \right) &\leq 2n \binom{n + \lfloor 2\gamma n \rfloor}{\lfloor 2\gamma n \rfloor} \exp \left( -\frac{9\beta c}{80R} n \right) \\ &\leq \exp \left( -\frac{9\beta c}{160R} n \right) \\ &\leq \frac{1}{n^3}. \end{aligned} \quad (13)$$

Writing  $\alpha = \min(\gamma, \epsilon, \delta)$ , it follows from (9), (10), and (13) that

$$\mathbf{P} \left( \exists S \in \bigcup_{R \log n \leq j \leq (1-\beta)n} B_j(H), e_H(S, S^c) \leq \alpha |S| \right) \leq \frac{3}{(n')^3}.$$

Since  $n' = n/R$  this completes the proof.  $\square$

## REFERENCES

- [1] B. Bollobás and F.R.K. Chung. The diameter of a cycle plus a random matching. *SIAM Journal on discrete mathematics*, 1:328–333, 1988.
- [2] G. Brightwell and P. Winkler. Maximum hitting time for random walks on graphs. *Random Structures & Algorithms*, 1(3):263–276, 1990.
- [3] S. Condamin, O. Bénichou, V. Tejedor, R. Voituriez, and J. Klafter. First-passage times in complex scale-invariant media. *Nature*, 450(7166):77–80, 2007.
- [4] R. Durrett. *Random graph dynamics*, volume 20. Cambridge University Press, 2007.
- [5] Illés J. Farkas, Imre Derényi, Albert-László Barabási, and Tamás Vicsek. Spectra of “real-world” graphs: Beyond the semicircle law. *Phys. Rev. E*, 64(2):026704, Jul 2001. doi: 10.1103/PhysRevE.64.026704.
- [6] N. Fountoulakis and B.A. Reed. Faster mixing and small bottlenecks. *Probability Theory and Related Fields*, 137(3):475–486, 2007.
- [7] L.K. Gallos, C. Song, S. Havlin, and H.A. Makse. Scaling theory of transport in complex biological networks. *Proceedings of the National Academy of Sciences*, 104(19):7746, 2007.
- [8] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley, New York, 2000.

- [9] F. Jasch and A. Blumen. Target problem on small-world networks. *Physical Review E*, 63(4):041108, 2001.
- [10] S. Jespersen, IM Sokolov, and A. Blumen. Relaxation properties of small-world networks. *Arxiv preprint cond-mat/0004214*, 2000.
- [11] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.
- [12] L. Lovasz and P. Winkler. Mixing Times. *Microsurveys in Discrete Probability: Dimacs Workshop, June 2-6, 1997*, 1998.
- [13] R. Montenegro and P. Tetali. Mathematical aspects of mixing times in markov chains. *Theoretical Computer Science*, 1(3):237–354, 2006.
- [14] M.E.J. Newman and DJ Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4-6):341–346, 1999.
- [15] M.E.J. Newman and D.J. Watts. Scaling and percolation in the small-world network model. *Physical Review E*, 60(6):7332–7342, 1999.
- [16] M.E.J. Newman, A.-L. Barabási, and D.J. Watts. *The structure and dynamics of networks*. Princeton Univ Pr, 2006.
- [17] J.D. Noh and H. Rieger. Random walks on complex networks. *Physical review letters*, 92(11):118701, 2004.
- [18] O. Riordan and N. Wormald. The diameter of sparse random graphs. *Combinatorics, Probability and Computing*, 19(5-6):835–926, 2010.
- [19] Richard P. Stanley. *Enumerative Combinatorics. Vol 2*. Cambridge University Press, Cambridge, 1999.
- [20] D.J. Watts and S.H. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393(6684):440–442, 1998.

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCGILL UNIVERSITY, 805 SHERBROOKE STREET WEST, MONTRÉAL, QUÉBEC, H3A 2K6, CANADA

*E-mail address:* `louigi@math.mcgill.ca`

*E-mail address:* `tao.lei@mail.mcgill.ca`